July 2013 Documentation Telecon

Attendees: Ted Haberman, Ed Armstrong, Curt Tilmes, Kelly Monteleone, Erin Robinson, Mike McCann, Bob Simon, Phil Jones, Lindsay, Walt Baskin

* Documentation had sessions at the ESIP summer meeting last week – they went well
  + 1st – lessons learned included GRIST, G6, CF & Data issues
  + 2nd – ACDD – need to get Dave’s presentation up as there were technical problems on webex.
    - Ted and Ed presented real world examples
  + Dave’s presentation suggested that documentation might move towards CF, but Ed though we had agreed to move to OGC standards
* ACDD – needs steering committee meeting to discuss name changes (suggested by John)
  + Minimized changes in names in ACDD 1.2 and address these in 2.0
* Bob – stresses the need for backwards compatibility
  + Avoid duplicating or replacing names
* Mike – focus on clarity/meaning of terms (need to work on definitions) because people come at this terms from different directions
  + Ex. contributor is principle investigator (PI) in ISO
  + Need better fit in ACDD for people
* Bob – ‘contributor role’ is gone from website
  + Ted – John added a “information” field and removed ‘contributor role’
* Mike – want to flatten into ACDD
  + Gulf of Mexico Research group has a tool for ISO metadata that has 3 ‘roles’
  + Ted – there are a large number of roles in ISO, ~20
  + In ACDD need 20 x 3 attribute names (name, email, url)
  + Ed – groups will help organize ‘role’ into discrete ‘objects’
  + ACDD 2.0 will have objects
* Mike – what about legacy data like NetCDF3
  + If flatten, all people can be put into ‘ResponsibleParty\_role\_name’ (or email, organization, etc).
  + To have organization in NetCDF3 – need all roles clarified
    - If the attribute is PI\_name then parser won’t know that this is a person
    - Also a problem if there are more than 1 PI or user support person
    - HDF has unique names
* Walt – attribute values in ACDD of a name can be an array of strings
  + - Ted – then have to match an array of strings across attributes (… email, position name, institution) – properties of people
    - (primary is 1st in list)
  + Ted – if information designed with delimiters in items – it causes problems with the design (Ed – it gets messy fast)
* Mike – there are real limitations in flat data models – want a simple flat model
  + New HDF conversion – when group node – can have multiple instances in group with same name
  + ‘groupname\_attribute’
* Ted – maybe ‘\_’ is a problem, could use “.” – Mike or “\_\_”…. These all cause other issues
* When is OceanSite going to use NetCDF4 (asked to Mike)?
  + It isn’t even on the radar – looking at NetCDF3
  + Will have to live with a flat data model
  + Strategy is to use a dumbed down ACDD that might spur a move to NetCDF4
* Bob – does this mean we need to either flatten ACDD or not support other file types – because groups are only in HDF5 or NetCDF4 extended.
  + do we need a list of other file types?
* ACDD 1.2 will be a flat model
* ACDD 2.0 will have objects (there will be a separate ACDD 2.0 that is flattened)
* Walt – if 1st had list of group (like a group parameter) that listed prefixes – then could clearly see different between groups and attributes
  + Even if 100s of groups – would be able to decode the structure
  + Like a map of how groups are put together
* Ted – in terms of flattening mechanisms – we need 3 – 4 to start testing (hopefully not 40-50)
  + Bob – need only 1 flattening method
  + Ted – yes, but need 4 or 5 so we can find the best
* Mike – need to ensure “lineage” information is not lost when move
  + Ted’s example last week had 9+ objects … all of them are needed
  + xml and Json have methods to parse
  + currently only have object (which is “group”), but need type for group – such as role, lineage, identities.
* Bob – has 3 issues
  + Want statement of “backwards compatibility” from this group
  + Limit versions of ACDD
  + Are we really just moving to HDF5 or will ACDD remain for file that do not include groups
* Mike – felt that ACDD 1.2 is gong backwards
* Ted – only Ted has commented on John’s discussion and ACDD edits
  + Bob – we need a system to review edits like CF so we can propose changes
  + Such as minicode changes are not added until formal agreement is reached… something like Jira (sp?)
  + Link to ACDD discussion on wiki - <http://wiki.esipfed.org/index.php/Talk:Attribute_Convention_for_Data_Discovery_(ACDD)_Working>
* On the ACDD page – the definitions say “an include any information as ISO 19139 or free text”. This is not helpful, need more information
* Currently the ‘watch” function on the wiki is not working… but recommend watching the ACDD discussion
* Next meeting
  + Rich
  + Maybe discussion – ACDD 1.2 finalize and definitions
  + Algorithums for flattening – may need wiki page
  + Other file formats – think about if we don’t need formal discussion of this
* Next meeting Aug 21 – noon EDT

|  |  |  |
| --- | --- | --- |
| **Description:** |  |  |
| **Streaming recording link:** |  | |  | | --- | | <https://esipfed.webex.com/esipfed/ldr.php?AT=pb&SP=MC&rID=69100652&rKey=d71045ca2da06e02> | |
| **Download recording link:** |  | |  | | --- | | <https://esipfed.webex.com/esipfed/lsr.php?AT=dw&SP=MC&rID=69100652&rKey=12bdca17e2495a04> | |